

(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau



(43) International Publication Date
26 June 2003 (26.06.2003)

PCT

(10) International Publication Number
WO 03/052590 A1

(51) International Patent Classification⁷: **G06F 9/00**

M.; 2005 Buckskin Drive, Clayton, NC 27520 (US).
ROBINSON, Robert, C.; 1613 Rogers Pointe Lane, Creedmoor, NC 27522 (US).

(21) International Application Number: PCT/US02/38353

(74) Agent: **ANDERSON, Jay, H.**; International Business Machines Corporation, Dept. 18G , Bldg. 300/482, 2070 Route 52, Hopewell Junction, NY 12533 (US).

(22) International Filing Date: 26 November 2002 (26.11.2002)

(81) Designated States (*national*): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VN, YU, ZA, ZM, ZW.

(25) Filing Language: English

(84) Designated States (*regional*): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),

(26) Publication Language: English

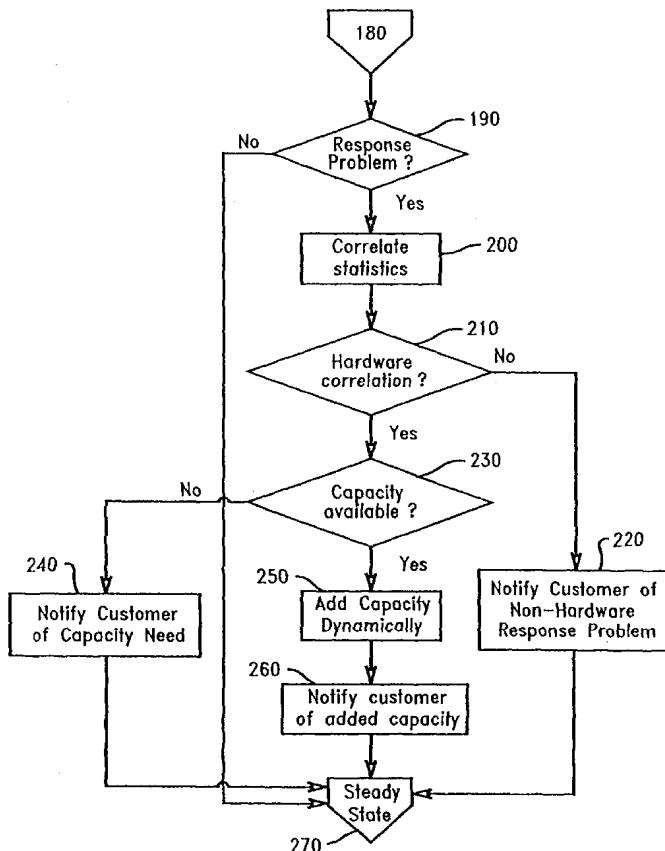
(30) Priority Data:
10/023,154 17 December 2001 (17.12.2001) US

(71) Applicant: **INTERNATIONAL BUSINESS MACHINES CORPORATION [US/US]**; Old Orchard Road, Armonk, NY 10504 (US).

(72) Inventors: **MOLLOY, Christopher, L.**; 1313 Hedgefern Way, Raleigh, NC 27615 (US). **QUESENBERRY, Gary,**

[Continued on next page]

(54) Title: AUTOMATIC DATA INTERPRETATION AND IMPLEMENTATION USING PERFORMANCE CAPACITY MANAGEMENT FRAMEWORK OVER MANY SERVERS



(57) Abstract: An automated method of managing computing resources having a workload of a given type comprises providing resource data collectors (110) for collecting data regarding performance of the resources, such as a server network, in accordance with the type of workload; developing a forecast (180) of utilization of the resources, based on historical performance data; and collecting real-time performance data (140) regarding the resources running under the workload. The method then includes analyzing the performance data and the forecast (190) to identify a critical resource (210) and automatically adjusting (250) a capacity of the resource to provide (270) steady-state performance of the resource under the workload. Additional hardware resources (40, 42, 60) such as CPUs, computer memory and computer disk storage are initially available to, but unused by, the computing resources.

WO 03/052590 A1



Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

Published:

— *with international search report*

**AUTOMATIC DATA INTERPRETATION AND IMPLEMENTATION
USING PERFORMANCE CAPACITY MANAGEMENT FRAMEWORK
OVER MANY SERVERS**

5

Technical Field

The present invention relates generally to computer systems and, in particular, to a method of adding critical hardware resource capacity to a networked system of computers.

10 **Background Art**

As described in U.S. Patent No. 6,148,335, the disclosure of which is incorporated herein by reference, a generalized client-server computing network has a plurality of servers and which are interconnected, either directly to each other or indirectly through one of the other servers. Each server is essentially a stand-alone computer system (having 15 one or more processors, memory devices, and communications devices), but has been adapted (programmed) for the primary purpose of providing information to individual users at a plurality of workstation clients in communication with each server. A client is a member of a class or group of computers or computer systems that uses the services of another class or group to which it is not related. As used herein, "client" generally refers to 20 any multi-purpose or limited-purpose computer adapted for use by a single individual, regardless of the manufacturer, hardware platform, operating system, and the like. The information provided by a server can be in the form of programs which run locally on a given client, or in the form of data such as files used by other programs.

Such networks may communicate via the Internet using conventional protocols and 25 services which allow the transfer of various types of information, including electronic mail, simple file transfers via FTP, remote computing via TELNET, gopher searching, Usenet newsgroups, and hypertext file delivery and multimedia streaming via the World Wide Web (WWW). A given server can be dedicated to performing one of these operations, or running multiple services. The '335 patent discloses the monitoring of server 30 performance in a network like the Internet, and generating reports detailing performance statistics (daily, weekly, or monthly) for various server resources. Statistical parameters may include for example the number of observations; CPU utilization; system usage percentage; user usage percentage; percentage of time I/O wait is greater than some pre-

selected level; run queue length; active virtual memory (AVM); free space (FRE); percentage of time CPU utilization is greater than some pre-selected level; percentage of time run queue is greater than some pre-selected level; percentage of time storage usage is greater than some pre-selected level and percentage of time paging rate is greater than 5 some number of pages per second. Links may be provided to view additional, detailed information regarding, for example, a specific resource on a particular server. Notwithstanding the advantages of the invention of the '335 patent, there is no method or system which may act on the performance information generated on the client-server computer network to improve the performance and reliability of the network.

10

Disclosure of Invention

Bearing in mind the problems and deficiencies of the prior art, it is therefore an object of the present invention to provide a method of monitoring and controlling network computer resources.

15

It is another object of the present invention to provide a method of managing computing resources on a network to improve the network's performance and reliability.

A further object of the invention is to provide a method by which critical hardware resource conditions may be addressed without direct user input.

20

Still other objects and advantages of the invention will in part be obvious and will in part be apparent from the specification.

The above and other objects and advantages, which will be apparent to one of skill in the art, are achieved in the present invention which is directed to, in a first aspect, an automated method of managing computing resources having a workload of a given type.

The method comprises providing resource data collectors for collecting data regarding

25

performance of the resources, in accordance with the type of workload; developing a forecast of utilization of the resources, based on historical performance data; and collecting real-time performance data regarding the resources running under the workload.

The method then includes analyzing the performance data and the forecast to identify a critical resource and automatically adjusting a capacity of the resource to provide steady-

30

state performance of the resource under the workload.

In another aspect, the present invention is directed to a program storage device readable by a machine, tangibly embodying a program of instructions executable by the machine to perform an automated method of managing computing resources having a workload of a given type, using resource data collectors for collecting data regarding

- performance of the resources in accordance with the type of workload, and a forecast of utilization of the resources based on historical performance data. The method steps comprise collecting real-time performance data regarding the resources running under the workload, analyzing the performance data and the forecast to identify a critical resource,
- 5 and automatically adjusting a capacity of the resource to provide steady-state performance of the resource under the workload.

In both of the above aspects of the invention, the resources preferably comprise a server network. The method may further comprise setting threshold values for the performance data and identifying the resource in accordance with the threshold values.

10 The method may also comprise notifying a user of the computing resources when the critical resource is a hardware resource, and notifying the user when the capacity of the hardware resource is adjusted. Preferably, the method further includes initially providing additional hardware resources available to, but unused by, the computing resources. Such additional hardware resources may be selected from the group consisting of CPUs,

15 computer memory and computer disk storage.

In yet another aspect, the present invention is directed to a computer program product for performing an automated method of managing computing resources having a workload of a given type, using resource data collectors for collecting data regarding performance of the resources in accordance with the type of workload, and a forecast of utilization of the resources based on historical performance data. The computer program product has computer-readable program code for collecting real-time performance data regarding the resources running under the workload, computer-readable program code for analyzing the performance data and the forecast to identify a critical resource, and computer-readable program code for automatically adjusting a capacity of the resource to provide steady-state performance of the resource under the workload.

As before, in this aspect the resources preferably comprise a server network. There may be initially provided additional hardware resources available to, but unused by, the computing resources. Such additional hardware resources may be selected from the group consisting of CPUs, computer memory and computer disk storage. The computer program product may further comprise computer-readable program code for setting threshold values for the performance data and computer-readable program code for identifying the resource in accordance with the threshold values. The computer program product may also comprise computer-readable program code for notifying a user of the computing resources when the critical resource is a hardware resource, and computer-readable

program code for notifying the user when the capacity of the hardware resource is adjusted.

Brief Description of Drawings

5 The features of the invention believed to be novel and the elements characteristic of the invention are set forth with particularity in the appended claims. The figures are for illustration purposes only and are not drawn to scale. The invention itself, however, both as to organization and method of operation, may best be understood by reference to the detailed description which follows taken in conjunction with the accompanying drawings
10 in which:

Fig. 1 is a flow chart of a portion of the preferred method of practicing the present invention.

Fig. 2 is a continuation of the flow chart of Fig. 1 showing the preferred method of practicing the present invention.

15 Fig. 3 is a schematic of one embodiment of a computer client/server network employing the method of the present invention.

Fig. 4 is a schematic of a preferred RISC computer having hardware resources which may be repartitioned among different partitioned servers.

20 Best Mode for Carrying Out the Invention

In describing the preferred embodiment of the present invention, reference will be made herein to Figs. 1-4 of the drawings in which like numerals refer to like features of the invention. Features of the invention are not necessarily shown to scale in the drawings.

25 The present invention is particularly useful in connection with the successful server resource management (SRM) methodology defined in the aforementioned U.S. Patent 6,148,335, whereby server resources are measured across multiple platforms and server trends reported by enterprise and/or server-level drill-down navigation using red/yellow/green report presentation. An online "red action list" of action plan and status is also reported. The automatic data interpretation of the present invention adds a layer of
30 benefit by implementing a set of automatic actions based on predefined correlation algorithms. When managing hundreds of installed machines, associated support costs are reduced through use of this management automation and alert methodology. In general, the present invention takes available server resource metrics for hardware resources such as central processing unit (CPU), memory and disk storage and develop framework to

automatically determine a set of actions based on measured conditions. This invention forms a closed loop whereby data is not only collected and reported at face-value, but also enables a set of recommendations or actions to be taken against the available data, saving analysis labor and intervention. The present invention provides a method to use capacity
5 on demand to add capacity automatically to the computer system, and to notify the user, e.g., the system manager or system analyst, when hardware capacity is added.

The present invention expands on the '335 patent and expands the list of actions to automatically recommend or implement capacity planning alternatives, such that the primary focus is server capacity planning. The present invention interprets server metrics
10 and workload resource data across platforms, and is not limited to mainframe data; it automatically determines a set of actions based on measured conditions, and uses statistical data and deduction techniques to perform the automation. The method and system of present invention are particularly directed to monitoring and analyzing server management data, as opposed to the business data on the server. The present invention
15 uses expected resources metric feeds and supports systems management of servers and/or Information Technology (I/T) machines; automatic interpretation is performed on the expected (server historical) data and rules set implemented; and uses relational database to archive the server history.

The method of the present invention may be described in detail in connection with
20 the flowcharts shown in Figs. 1 and 2. Initially, in step 110, server resource collectors are installed in the server system to collect data regarding performance of the server resources and threshold values for the performance data are defined. In subsequent steps, the system begins collecting and logging steady state server metrics by first starting the data collection process 120, determining the running workload of the server system 130,
25 starting the data collection for each workload 140, and setting the collector threshold based on the workload mix 150. The workloads and workload mix may be any combination of system and/or application processes such as web hosting, database hosting, file serving, security checking, batch processing, financial systems, network management, systems management, numerical and statistical analysis, online processing, and the like.
30 Using the historical (as opposed to real-time) steady state server data measurements as input 160, the method then develops (in steps 170, 180) metrics associated with CPU usage and thresholds to determine the need for additional CPU capacity.

Based on the information previously collected and the forecast computed, the method then determines whether there are any response or resource bottlenecks 190. These

are determined using specific platform metrics, such as page rate, run queue, scan rate, out-and-ready, swap rate; I/O rates, disk utilization, and the like. The server system response time is measured and statistics are correlated 200 to determine the threshold values to be set for use of the hardware resources. For example, disk storage capacity 5 threshold values may be set at some percentage of available disk space, or CPU usage may be set at some percentage of maximum usage. If threshold exceptions are found, then the method determines whether hardware resources are an impact 210. If no hardware contention correlation is found, then the customer is notified of response time threshold exceptions 220, with no hardware issues detected. If the critical resource capacity is 10 available on-demand 230, then such capacity is adjusted, i.e., added dynamically to the server 250, and the customer is notified of action taken 260. If a hardware resource contention is detected, i.e., a critical resource, and no additional hardware capacity is available, then the customer is notified of need for capacity 240. If no response time or hardware resource exceptions are detected, then regular steady state conditions 270 15 continue and the process continues again: (a) log date; (b) analyze and correlate data; (c) activate automation policy or alerting, as necessary; (d) notify customer of actions or resource status; and (e) continue steady state.

The critical hardware resource capacity added in step 250 may be any central processing unit (CPU) microprocessor, computer memory, storage, or other hardware 20 resource necessary to maintain the system at steady state operation. Memory devices may include random access memory (RAM), read only memory (ROM), and nonvolatile memory (e.g., EPROM, flash memory, or battery-pack CMOS RAM). Storage includes disk such as optical (e.g., CD-ROM) or magnetic drives, or other storage media. Fig. 3 depicts one possible hardware arrangement for use with the present invention. Servers 10, 25 12, 14, 16 are linked to each other as a plurality of network nodes operating on the same or different platforms. Each server is linked to a plurality of client computers, i.e., server 10 is linked to clients 20a, 20b, 20c, server 12 is linked to clients 22a, 22b, 22c, server 14 is linked to clients 24a, 24b, 24c, and server 16 is linked to clients 26a, 26b, 26c. The clients 30 may be stand-alone personal computers or limited-use network computers. The links between the various clients and servers are sufficient to transfer the types of information used on the particular network, such as the aforedescribed Internet protocols and services.

The server resource management (SRM) architecture collects data using a remote command facility (RCF) program on server 30 which works by executing UNIX commands to gather utilization data from one or more servers, such as by the scripting

language known as PERL (practical extraction and report language) to issue the commands which gather the bulk of the data. The UNIX or other machine- or computer-readable program code used by the RCF may be stored on any of the storage media described above. The RCF process can use the low- impact "sockets" interface, and be 5 extensible for executing data gathering commands on other brands of UNIX. RCF collects key server resource data including current CPU utilization, memory availability, I/O usage, and permanent storage (disk) capacity. An output file is generated containing the collected information, which can be stored locally on a hard disk drive or at a remote location, preferably not one of the servers being monitored. The RCF can provide a user 10 interface for data collection by using conventional communications software such as a web browser that is adapted to display a page having commands or tool bars used to manage data collection. Other communications software can be used besides standard web browsers, such as those described in the '335 patent. If a server does not run a UNIX-type platform, other commands can be used to collect the data, such as those also described in 15 the '335 patent. In the foregoing manner, key performance and capacity data from a wide variety of servers becomes web-accessible. Data collection from different servers can occur at different times, i.e., there is no need for data processing system on server 30 to be continuously connected to each of the servers.

Once the data has been collected, it can be deposited into an appropriate database 20 as described, and optionally merged with other historical data previously collected. The collected data can then be forwarded (e.g., via FTP) to a node running an analysis program, such as the Statistical Analysis System (SAS). This software provides a programming language used to analyze data processing. The analyzed data can be presented in a variety of media or formats. In one implementation, a web browser can 25 again be used to view the analysis, by creating an HTML file which is then placed on the network (e.g., the World Wide Web) in such a manner as to be accessible and usable by the end-user.

Going beyond merely reporting server performance, the present invention has the ability to adjust automatically capacity of the hardware resources identified as deficient by 30 the SRM. The method of the present invention, as described in Figs. 1 and 2, is programmed by conventional programming code in the RCF program. As shown in Fig. 3, the additional hardware resources, such as CPU, memory or disk storage 40, 42, is available to be linked to the server on command of the RCF. These additional hardware resources may be available to a plurality of servers, such as resource 42 is available to

servers 12, 14, 16, or the additional hardware resources may be available to a single server, such as resource 42 is available only to servers 10. The links from the additional hardware resources to the servers are activated and connected on command by the RCF when added capacity is determined by the RCF to be needed by one or more of the 5 servers. Thus, the method and system of present invention are able to analyze not only the utilization of the entirety of the server install base, but also exceptions at individual, single servers which require additional resources.

Industrial Applicability

10 A preferred computer system on which to use the method of the present invention is a mid-level computer partitioned to operate as a plurality of separate servers, and capable of being re-partitioned to reallocate critical hardware resources among the different servers. This system is depicted in schematic in Fig. 4 wherein the total computer system hardware and software resources 50 are subdivided into subsystems, here showing 15 some of such subsystems as virtual, separate servers 10a, 10b, 10c, 10d, 10e, 10f, 10g and 10h, each linked to each other for conventional network communication. The desired hardware resources which are available to be dedicated to or shared among the virtual servers, such as CPU, memory and disk storage, are shown as 60. Initially, a predetermined amount of the CPU, memory, disk storage and other hardware resource is 20 dedicated to each virtual server. As the network operates normally under its workload, the RCF determines whether additional resource capacity is needed at one or more of the virtual servers, in accordance with the method of the present invention. The RCF then repartitions available resources 60 to add the identified critical resource to the virtual server requiring it, to restore steady state performance. A computer system having such 25 partitionable hardware resources is under development as a RISC 6000 system from IBM Corporation, Armonk, NY.

The present invention automatically determines what workload is running on the computer, starts collectors based on type of workload, sets thresholds for metrics based on workload mix, determines when metrics exceed threshold (both current and projected 30 workload), and correlates metrics to determine if hardware capacity is the cause of the problem. Additionally, the present invention automatically correlates server metrics with available middleware metrics to enable problem detection. By this method, it is possible to determine automatically if extra capacity exists, determine resource bottlenecks using

historical data, add capacity if available, automatically notify people of actions taken, and provide a customer interface to set custom resource thresholds.

While the present invention has been particularly described in conjunction with a specific preferred embodiment, it is evident that many alternatives, modifications and variations will be apparent to those skilled in the art in light of the foregoing description. It is therefore contemplated that the appended claims will embrace any such alternatives, modifications and variations as falling within the true scope and spirit of the present invention.

10

Claims

- 1 1. An automated method of managing computing resources having a workload of
2 a given type, the method comprising the steps of:
3 providing resource data collectors (110) for collecting data regarding performance
4 of the resources, in accordance with the type of workload;
5 developing a forecast (180) of utilization of the resources, based on historical
6 performance data (160);
7 collecting real-time performance data (140) regarding the resources running under
8 the workload;
9 analyzing the real-time performance data and the forecast (190) to identify a
10 critical resource (210); and
11 automatically adjusting (250) a capacity of the resource to provide (270) steady-
12 state performance of said resource under said workload.

- 1 2. The method of claim 1, characterized in that said resources comprise a server
2 network (10).

- 1 3. The method of claim 1 further comprising setting threshold values for said
2 performance data (200) and identifying the resource in accordance with the
3 threshold values.

- 1 4. The method of claim 1 further comprising the steps of:
2 notifying a user of the computing resources when the critical resource is a
3 hardware resource; and
4 notifying the user when the capacity of said hardware resource is adjusted (260).

- 1 5. The method of claim 1 further comprising initially providing additional
2 hardware resources (40,42,60) available to, but unused by, the computing
3 resources.

- 1 6. The method of claim 5, characterized in that the additional hardware resources
2 are selected from the group consisting of CPUs, computer memory and
3 computer disk storage.

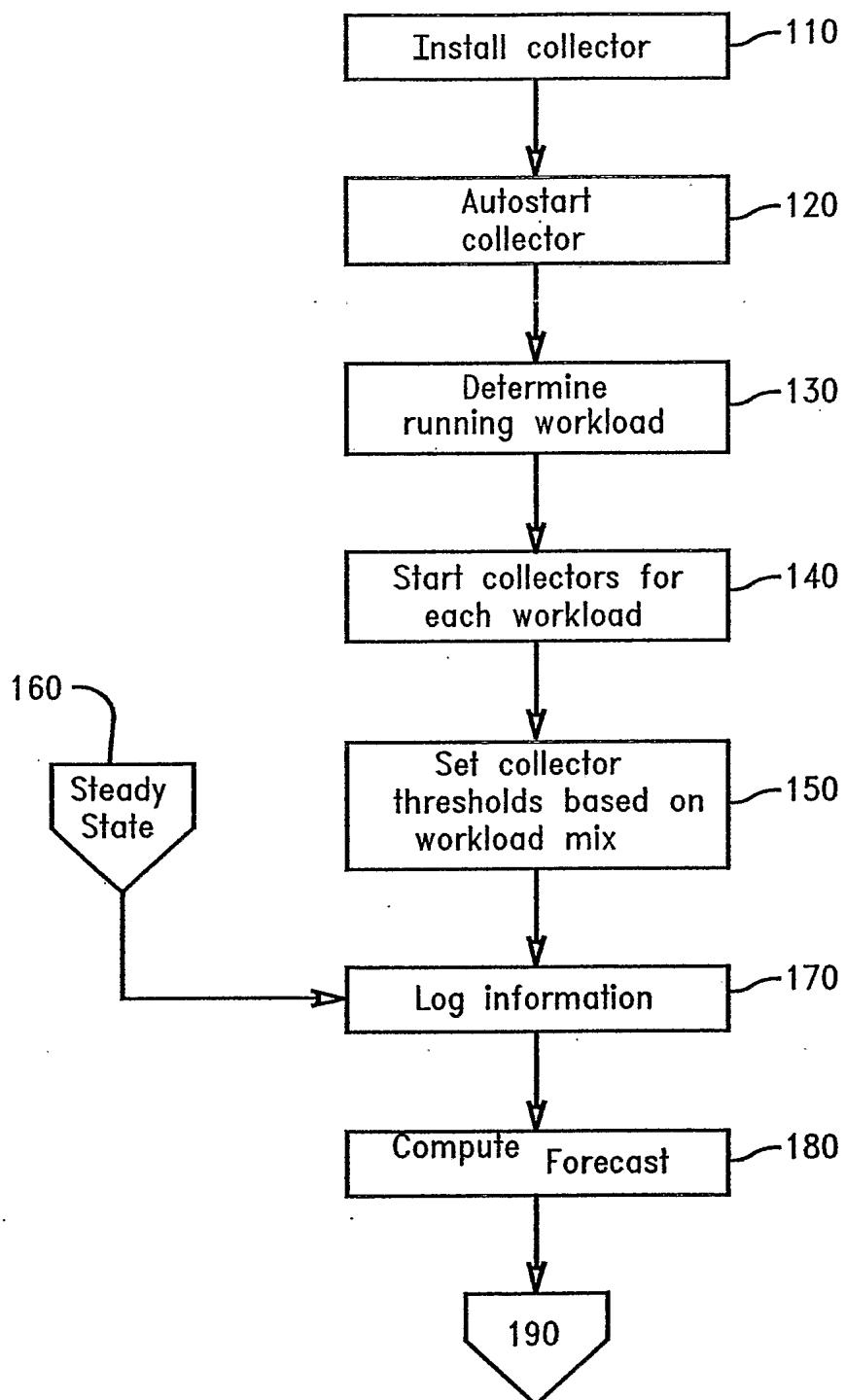


FIG. 1

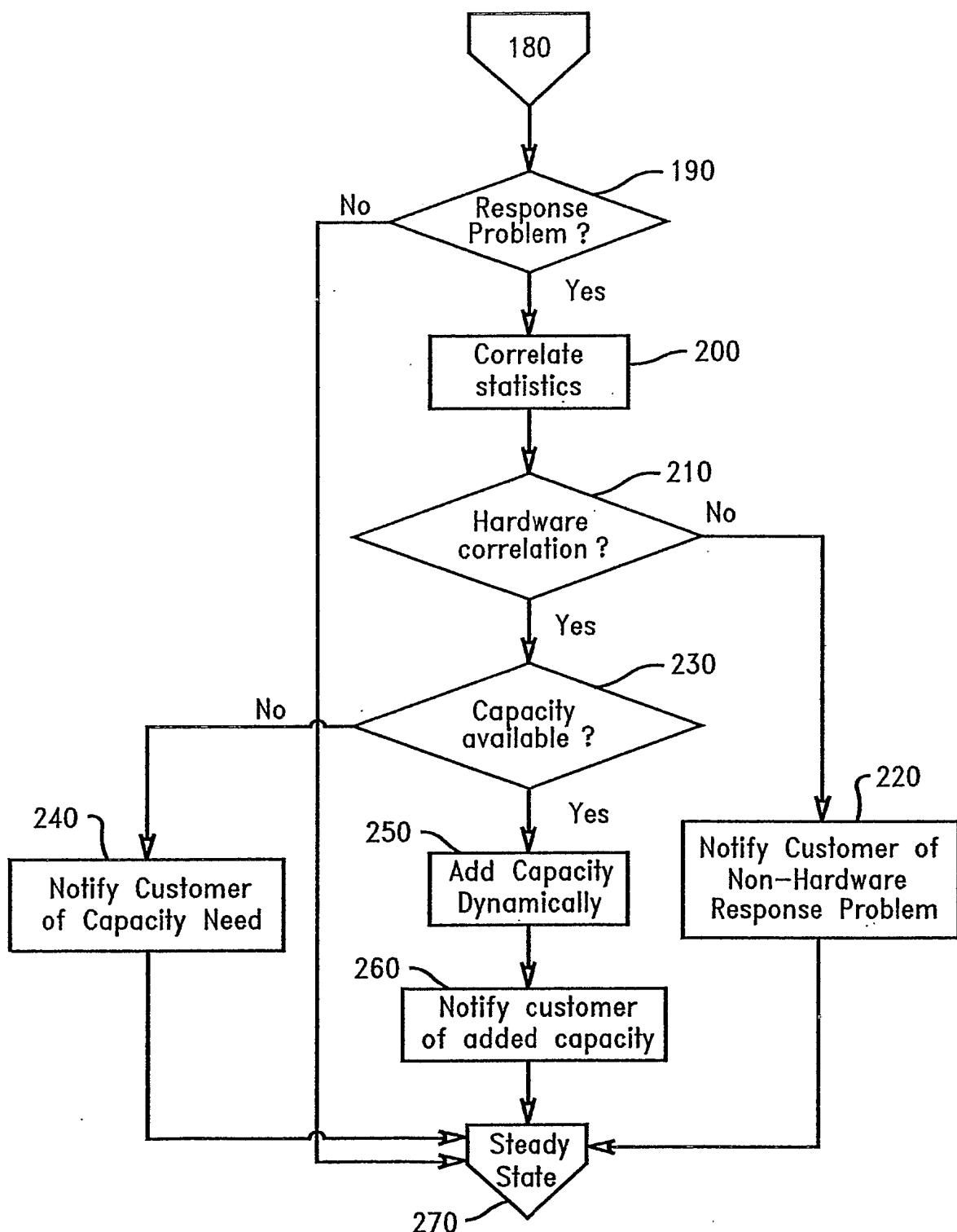


FIG. 2

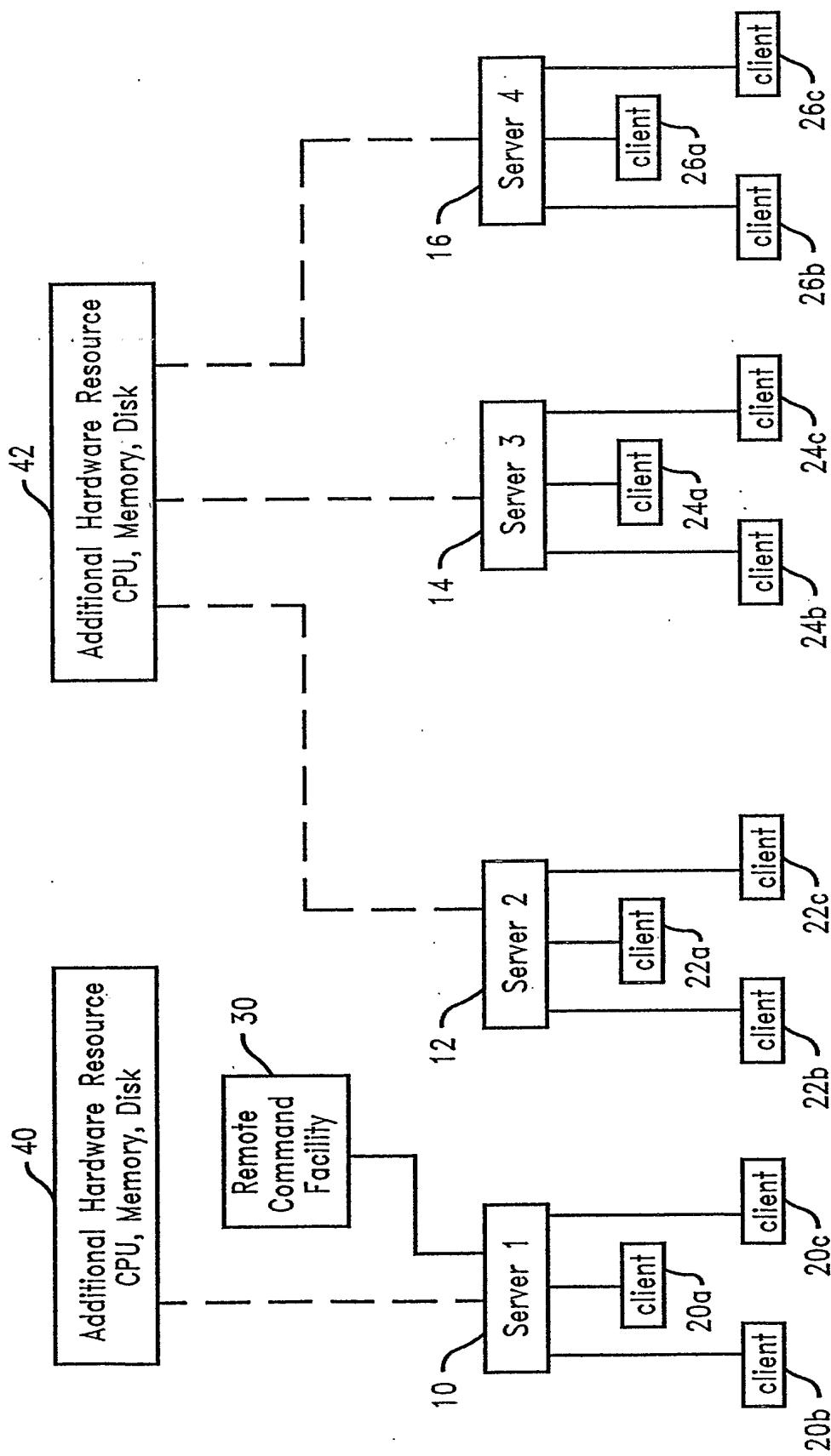


FIG. 3

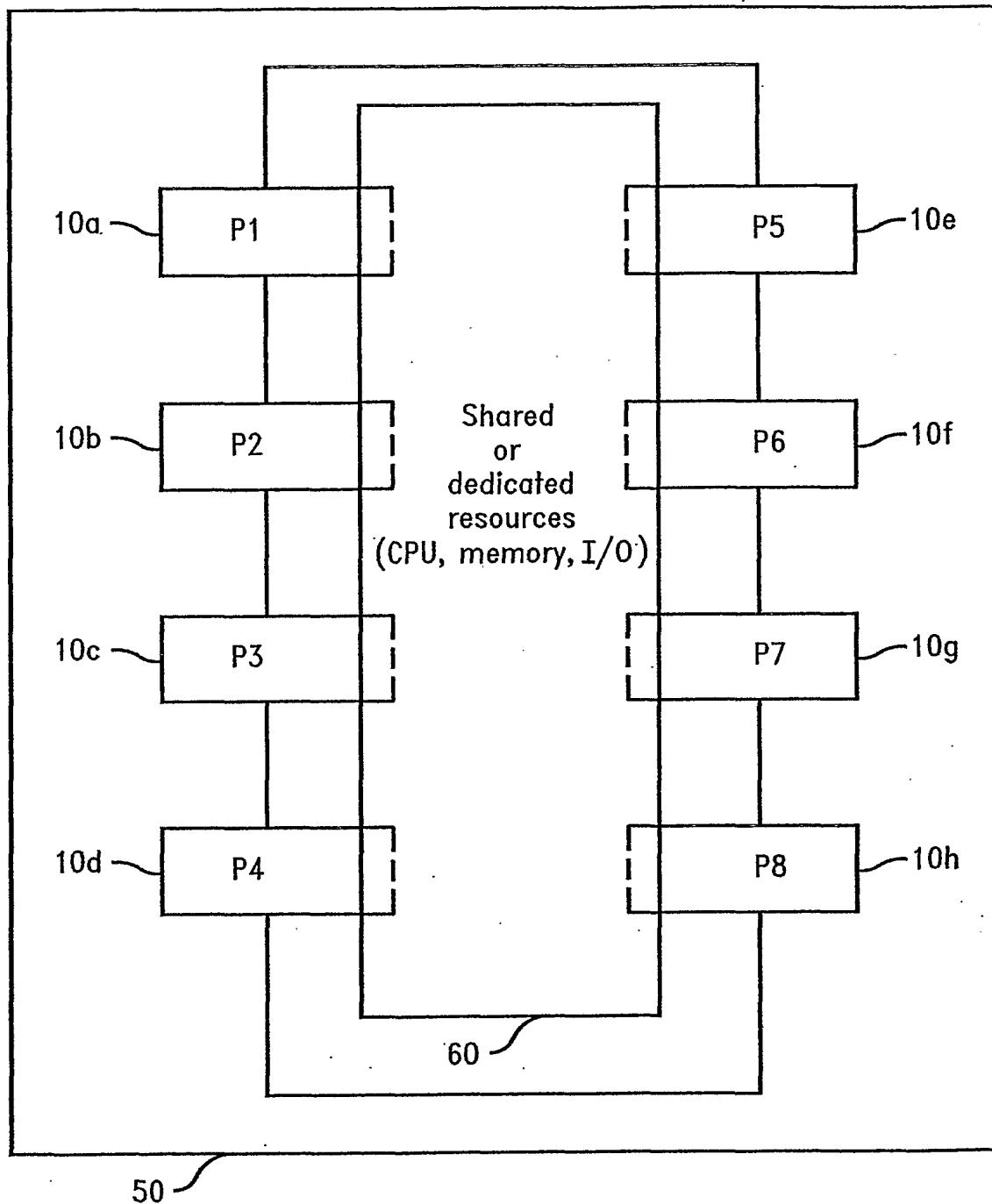


FIG. 4

INTERNATIONAL SEARCH REPORT

International application No.

PCT/US02/38353

A. CLASSIFICATION OF SUBJECT MATTER

IPC(7) : G06F 9/00
US CL : 709/100

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
U.S. : 709/100, 709/200, G06F 15/16

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EAST

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X, P	US Pat. Application Publication 2002/0116441 A1 (DING ET AL.) 22 Aug 2002, fig. 1, and 3 - 4, Paragraph 0035, 0157, 0123, 033, and 0074.	1 - 3 and 5 - 6
Y, P	US 6,453,353 B1 (WIN ET AL.), 17 September 2002, col. 8, lines 47 - 63, col. 28, lines 19 - 27.	4

Further documents are listed in the continuation of Box C.

See patent family annex.

*	Special categories of cited documents:	"T"	later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"A"	document defining the general state of the art which is not considered to be of particular relevance	"X"	document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"E"	earlier application or patent published on or after the international filing date	"Y"	document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"L"	document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"&"	document member of the same patent family
"O"	document referring to an oral disclosure, use, exhibition or other means		
"P"	document published prior to the international filing date but later than the priority date claimed		

Date of the actual completion of the international search

22 January 2003 (22.01.2003)

Date of mailing of the international search report

10 FEB 2003

Name and mailing address of the ISA/US

Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. (703)305-3230

Authorized officer

John Follansbee

Telephone No. 703-305-3900